Problem 1.2

Let PV stands for percent vectorization divided by 100.

a. Plot Net speedup overall = $1 / \left( (1 - F_{enh}) + \frac{F_{enh}}{S_{up}} \right)$, for $0 \leq F_{enh} \leq 1$.

b. From equation in (a). If Net speedup = 2 then the percent vectorization is $5/9$ or $56\%$.

c. Time in vector mode = $PV / (10 \times (1 - PV + PV / 10)) = 1/9$ or $11\%$. 
d. From the equation in (a). If Net speedup = 10/2 = 5 then PV = 8/9 or 89%.

e. The increased percent vectorization needed to match a hardware speedup of $10 \times 2 = 20$ applied to the original 70% vectorization is

$$1 / (1 – PV + PV / 10) = 1 / (1 – 70\% + 70\% / 20)$$

Solving shows that vectorization must increase to 74%, not a large increase. Improving the compiler to increase vectorization another 4% may be easier and cheaper than improving the hardware by a factor of 2.

Problem 1.3

a) Amdahl’s law states that the speedup is given by

$$\text{Speedup}_{\text{overall}} = \frac{\text{Time}_{\text{unchanged}}}{\text{Time}_{\text{enhanced}}}$$

The unenhanced time is the sum of the time that does not benefit from 10 times faster speed up + time that does benefit, but before the reduction by the factor

$$\text{Time}_{\text{unchanged}} = 50\% \times \text{Time}_{\text{enhanced}} + 10 \times 50\% \times \text{Time}_{\text{enhanced}}$$

$$= 5.5 \times \text{Time}_{\text{enhanced}}$$

Substituting in the equation of Speedup we get

$$\text{Speedup}_{\text{overall}} = \frac{(5.5 \times \text{Time}_{\text{enhanced}})}{(\text{Time}_{\text{enhanced}})} = 5.5$$

Thus,

**Speedup = 5.5**
b) According to Amdahl’s Law we have

\[
\text{Speedup}_{\text{overall}} = \frac{1}{(1 - \text{Fraction}_{\text{enhanced}}) + \text{Fraction}_{\text{enhanced}}} \cdot \frac{\text{Speedup}_{\text{enhanced}}}{\text{Speedup}_{\text{enhanced}}}
\]

Here it is given that,

\[
\text{Speedup}_{\text{enhanced}} = 10.
\]

Substituting in the above equation we get,

\[
5.5 = 1 / (1 - \text{Fraction}_{\text{enhanced}}) + (\text{Fraction}_{\text{enhanced}}/10)
\]

Solving the above equation we get,

\[
\text{Fraction}_{\text{enhanced}} = 0.909 = 91\%
\]

**Problem 1.8**

a)

\[
\text{Dies per wafer} = \frac{\pi \times (\text{wafer diameter}/2)^2}{\text{Die area}} - \frac{\pi \times \text{wafer diameter}}{\sqrt{2} \times \text{Die area}}
\]  \hspace{1cm} (1)

\[
\text{Die yield} = \text{wafer yeild} \times (1 + \frac{\text{defects per unit area} \times \text{Die area}}{\alpha})^{-\alpha}
\]  \hspace{1cm} (2)

\[
\text{Number of good chips} = \text{Dies per wafer} \times \text{Die yield}
\]

Use Alpha21264C as example:

\[
\text{Dies per wafer} = \frac{\pi \times (200/2)^2}{115} - \frac{\pi \times 200}{\sqrt{2} \times 115} = 273 - 42 = 231
\]

\[
\text{Die yield} = 0.95 \times (1 + \frac{0.5 \times 1.15}{4})^{-4} = 0.555
\]
Number of good chips = 231 × 0.555 = 128

Use the same formula for the other microprocessors and get the following:

<table>
<thead>
<tr>
<th>Microprocessor</th>
<th>Dies/wafer</th>
<th>Good dies/wafer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha 21264C</td>
<td>231</td>
<td>128</td>
</tr>
<tr>
<td>Power3-II</td>
<td>157</td>
<td>71</td>
</tr>
<tr>
<td>Itanium</td>
<td>79</td>
<td>20</td>
</tr>
<tr>
<td>MIPS R14000</td>
<td>122</td>
<td>46</td>
</tr>
<tr>
<td>UltraSPARC III</td>
<td>118</td>
<td>44</td>
</tr>
</tbody>
</table>

b)  
\[ \text{Cost of die} = \text{Cost of wafer} \times \frac{\text{Die yield}}{\text{Dies of wafer} \times \text{Die yield}} = \frac{\text{Cost of wafer}}{\text{Number of good chips}} \]

Use Alpha21264C as example:

\[ \text{Cost of die} = \frac{4700}{128} = 36.72 \]

Use the same formula for the other microprocessors and get the following:

<table>
<thead>
<tr>
<th>Microprocessor</th>
<th>$/good die</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha 21264C</td>
<td>$36.72</td>
</tr>
<tr>
<td>Power3-II</td>
<td>$56.34</td>
</tr>
<tr>
<td>Itanium</td>
<td>$245.00</td>
</tr>
<tr>
<td>MIPS R14000</td>
<td>$80.43</td>
</tr>
<tr>
<td>UltraSPARC III</td>
<td>$118.18</td>
</tr>
</tbody>
</table>

c)  
\[ \text{Test Cost} = \frac{\text{Test cost per hour} \times \text{Test time(secs)}}{3600 \times \text{Die yield}} \]

Cost of good tested packaged die = Cost of die + test cost + package cost

Use Alpha21264C as example:

\[ \text{Cost of good tested packaged die} = \text{Cost of die} + \frac{440 \times 25}{3600 \times 0.555} + 25 = 67.22 \]
Use the same formula for the other microprocessors and get the following:

<table>
<thead>
<tr>
<th>Microprocessor</th>
<th>$/good,tested,packaged die</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha 21264C</td>
<td>$67.22</td>
</tr>
<tr>
<td>Power3-II</td>
<td>$81.50</td>
</tr>
<tr>
<td>Itanium</td>
<td>$277.53</td>
</tr>
<tr>
<td>MIPS R14000</td>
<td>$113.41</td>
</tr>
<tr>
<td>UltraSPARC III</td>
<td>$158.88</td>
</tr>
</tbody>
</table>

d) The largest processor in Figure 1.34 is Itanium:

Just follow the steps described in a), b) and c) except using different “defect per unit area” in formula (2). The answer is as follow:

<table>
<thead>
<tr>
<th>Itanium</th>
<th>$/good,tested,packaged die</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defect density=0.5</td>
<td>$277.53</td>
</tr>
<tr>
<td>Defect density=0.3</td>
<td>$176.38</td>
</tr>
<tr>
<td>Defect density=1.0</td>
<td>$665.50</td>
</tr>
</tbody>
</table>

e) Digital 21064A:

Die area: 166mm²
Pins: 431
Technology: CMOS 0.5micron 4.5M
Estimated Wafer Cost: 4000
Package:Ceramic PGA.

For a=4

\[
Dies\ per\ wafer = \frac{\pi \times (200/2)^2}{166} - \frac{\pi \times 200}{\sqrt{2} \times 166} = 189 - 35 = 154
\]

\[
Die\ yield = 0.95 \times (1 + \frac{0.8 \times 1.66}{4})^{-1} = 0.30
\]

\[
Number\ of\ good\ chips = 154 \times 0.30 = 46
\]
Cost of die \( \frac{\text{Cost of wafer}}{\text{Number of good chips}} = \frac{4000}{46} = 86.96 \)

Test Cost \( \frac{\text{Test cost per hour} \times \text{Test time (sec)}}{3600 \times \text{Die yield}} = \frac{400 \times 20}{3600 \times 0.30} = 7.4 \)

Cost of good tested packaged die = Cost of die + Test Cost + Package cost = 86.96 + 7.4 + 20 = 114.36

For a=6, the same:

Cost = 118.68

Some of you use the data of Alpha 21264C:

<table>
<thead>
<tr>
<th></th>
<th>Dies/wafer</th>
<th>Die yield</th>
<th>Good die</th>
<th>Cost of Die</th>
<th>Packet cost</th>
<th>Test cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>a=4</td>
<td>231</td>
<td>0.42</td>
<td>97</td>
<td>48.45</td>
<td>25</td>
<td>7.28</td>
<td>80.73</td>
</tr>
<tr>
<td>a=6</td>
<td>231</td>
<td>0.40</td>
<td>92</td>
<td>51.09</td>
<td>25</td>
<td>7.64</td>
<td>83.73</td>
</tr>
</tbody>
</table>

**Problem 1.17**

a) Using the symbols in the question the equations for the MIPS ratings for both the processor and processor/coprocessor configurations are as follows:

\[
\text{MIPS} = \frac{\text{Instruction Count}}{\text{Execution Time} \times 10^6}
\]

\[
\text{MIPS}_{\text{processor}} = \frac{I + F.Y}{W}
\]

\[
\text{MIPS}_{\text{processor/coprocessor}} = \frac{I + F}{B}
\]
b) We have
\[
\text{MIPS}_{\text{processor}} = \frac{I + F \cdot Y}{W}
\]

Now we have \( F = 8 \times 10^6 \)
\( Y = 50, \) and
\( W = 4 \)

Also we have MIPS w/o coprocessor as \( 120 \times 10^6 \)

Substituting values in the above equation we get
\[
I = 120 \times 10^6 (W) - 8 \times 10^6 (50)
= 80 \times 10^6 \text{ instructions}
\]

Thus we get,
\[
I = 80 \times 10^6 \text{ instructions}
\]

c) We have
\[
\text{MIPS}_{\text{processor/coprocessor}} = \frac{I + F}{B}
\]

Now we have \( F = 8 \times 10^6 \)
\( I = 80 \times 10^6 \text{ instructions} \)

Also we have
\[
\text{MIPS}_{\text{processor/coprocessor}} = 80 \times 10^6
\]

Substituting values in the above equation we get,
\[
B = \frac{80 \times 10^6 + 8 \times 10^6}{80 \times 10^6} \text{ seconds}
= 1.1 \text{ second}
\]

d) Now we have
\[
\text{MFLOPS} = \frac{\text{No. Of Floating Point Operations in a program}}{\text{Execution time (sec) for Floating Pt. Inst.} \times 10^6}
\]
\[
F = \frac{10^6}{B \times 10^6} - \text{Time for Integer Instructions}
\]

Now,
\[
\text{Time for Integer Instructions} = \frac{I}{\text{MIPS}_{\text{processor}}}
\]

Note: Here we are using \( \frac{I}{\text{MIPS}_{\text{processor}}} \) because with the processor alone all the instructions are executed as integer instructions. Even the floating point instructions are emulated as integer inst. (i.e. executed as integer instructions). So we can get the time for execution of integer instructions using the MIPS of the machine with processor only. Now when the coprocessor is added it only affects the execution of FP instructions.

\[
\frac{80 \times 10^6}{120 \times 10^6} = 0.66
\]

Thus we get
\[
\text{MFLOPS} = \frac{8 \times 10^6}{1.1 - 0.66}
\]

\[
= 18 \text{ MFLOPS}
\]

Thus we get,
\[
\text{MFLOPS} = 18
\]

e) The time for the processor alone is \( W = 4 \text{ sec} \) while that for the processor/coprocessor is \( B = 1.1 \text{ sec} \). This shows that the execution time of the processor/coprocessor is faster even though the MIPS rating is low. Your colleague’s evaluation is therefore correct.